Report on Overfitting and Underfitting with Regularization on CIFAR-10

Introduction

This report analyses the effects of different regularization techniques (L1 and L2) on the performance of a neural network trained on the CIFAR-10 dataset. The CIFAR-10 dataset consists of 60,000 32x32 colour images across 10 classes, making it a challenging task for image classification. The goal is to observe the training and evaluation accuracy and loss for various regularization strengths and to identify signs of overfitting and underfitting.

Results Summary

The following table summarizes the training accuracy, evaluation accuracy, training loss, and evaluation loss for different regularization techniques and strengths:

Regularization	Training Accuracy	Evaluation Accuracy	Training Loss	Evaluation Loss
L1 (0.001)	0.47228	0.4623	1.616889	1.626648
L1 (0.01)	0.09966	0.1000	2.452327	2.452203
L1 (0.1)	0.09934	0.1000	3.812289	3.806581
L2 (0.001)	0.70406	0.6712	1.011299	1.116748
L2 (0.01)	0.55208	0.5669	1.506753	1.471551
L2 (0.1)	0.09612	0.1000	2.302762	2.302596

Observations:

1. **L1 Regularization:**
    - **L1_0.001:**
        - **Training Accuracy:** 47.23%
        - **Evaluation Accuracy:** 46.23%
        - **Interpretation:** The model is underfitting, as indicated by the low accuracy and relatively high training and evaluation loss values. The regularization strength is too strong, penalizing the model's ability to learn effectively.
    - **L1_0.01 & L1_0.1:**

        - **Training Accuracy:** ~9.9%
        - **Evaluation Accuracy:** ~10.0%
        - **Interpretation:** Severe underfitting is observed with these higher L1 strengths. The model fails to learn, as evidenced by the minimal difference between training and evaluation accuracy, which hovers around chance level.

2. **L2 Regularization:**
    - **L2_0.001:**
        - **Training Accuracy:** 70.41%
        - **Evaluation Accuracy:** 67.12%
        - **Interpretation:** This configuration achieves the best balance, indicating appropriate regularization that reduces overfitting while maintaining model complexity.
    - **L2_0.01:**
        - **Training Accuracy:** 55.21%
        - **Evaluation Accuracy:** 56.69%
        - **Interpretation:** The model starts to underfit as the regularization strength increases, with a noticeable drop in both training and evaluation accuracy.
    - **L2_0.1:**
        - **Training Accuracy:** 9.61%
        - **Evaluation Accuracy:** 10.00%
        - **Interpretation:** Severe underfitting is evident, as the model's performance is significantly hindered by the high regularization strength.

Justification for Chosen Regularization Strength

Based on the observations, the L2 regularization with a strength of **0.001** appears to be the most effective choice for the CIFAR-10 dataset. This configuration achieves the highest training accuracy (0.70406) and a reasonably high evaluation accuracy (0.6712), indicating a good balance between fitting the training data and generalizing to unseen data.

- **Avoiding Overfitting**: The relatively low training and evaluation losses suggest that the model is not overfitting significantly, which is crucial for maintaining performance on unseen data.

- **Effective Learning**: The model with L2 (0.001) regularization is able to learn meaningful patterns from the data without being overly constrained by regularization, which is evident from its performance metrics.
- **Excessive Regularization**: Higher regularization strengths (L1 and L2 at 0.01 and 0.1) led to underfitting, as indicated by the low accuracies and high losses. This suggests that the model was too restricted to learn effectively from the training data.

Conclusion

In conclusion, the analysis of the CIFAR-10 dataset reveals that regularization plays a critical role in managing overfitting and underfitting. The L2 regularization with a strength of **0.001** provides the best balance, yielding the highest accuracy while minimizing loss, thus allowing the model to generalize effectively. Further experiments with different regularization techniques or hyperparameters may be beneficial to optimize performance further.